frontiers in Applied Mathematics and Statistics

METHODS
published: 25 January 2021
doi: 10.3389/fams.2020.611878

A New Nonparametric Estimate of the Risk-Neutral Density with Applications to Variance Swaps

Liyuan Jiang[1], Shuang Zhou[1], Keren Li[1], Fangfang Wang[2] and Jie Yang[1]*

[1]Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL, United States,
[2]Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA, United States

Estimates of risk-neutral densities of future asset returns have been commonly used for pricing new financial derivatives, detecting profitable opportunities, and measuring central bank policy impacts. We develop a new nonparametric approach for estimating the risk-neutral density of asset prices and reformulate its estimation into a double-constrained optimization problem. We evaluate our approach using the S&P 500 market option prices from 1996 to 2015. A comprehensive cross-validation study shows that our approach outperforms the existing nonparametric quartic B-spline and cubic spline methods, as well as the parametric method based on the normal inverse Gaussian distribution. As an application, we use the proposed density estimator to price long-term variance swaps, and the model-implied prices match reasonably well with those of the variance future downloaded from the Chicago Board Options Exchange website.

Keywords: pricing, risk-neutral density, double-constrained optimization, normal inverse Gaussian distribution, variance swap

OPEN ACCESS

Edited by:
Eleftherios Ioannis Thalassinos, University of Piraeus, Greece

Reviewed by:
Ramona Rupeika-Apoga, University of Latvia, Latvia
Inna Romānova, University of Latvia, Latvia

****Correspondence:***
Jie Yang
jyang06@uic.edu

Specialty section:
This article was submitted to Mathematical Finance, a section of the journal Frontiers in Applied Mathematics and Statistics

Received: *29 September 2020*
Accepted: *03 December 2020*
Published: *25 January 2021*

Citation:
Jiang L, Zhou S, Li K, Wang F and Yang J (2021) A New Nonparametric Estimate of the Risk-Neutral Density with Applications to Variance Swaps. Front. Appl. Math. Stat. 6:611878. doi: 10.3389/fams.2020.611878

1 INTRODUCTION

A financial derivative, such as option, swap, future, or forward contract, is an asset that is contingent on an underlying asset. Its fair price can be obtained by calculating the expected future payoff under a risk-neutral probability distribution. Therefore, the problem of pricing a derivative can be addressed via estimating the risk-neutral density (RND) of the future payoff of the underlying asset. The estimated prices, known as fair prices, may help companies and investors to avoid financial risk or detect profitable opportunities, especially on over-the-counter (OTC) securities. The estimated risk-neutral distribution can be further used by, for example, central banks to infer market belief on economic events of interest and measure impacts of monetary policies [1–3]. On the other hand, the market prices of the derivatives traded in a financial market reveal information about the RND. Breeden and Litzenberger (1978) [4] were among the first to use option prices to estimate the risk-neural probability distribution of the future payoff of the underlying asset. Among the financial products that can be used to recover the RND, European options are the most common ones, which give the investors rights to trade assets at a preagreed price (i.e., strike price) at the maturity date. Among all the underlying assets that options are written on, Standard and Poor's 500 Index (S&P 500) is a popular one, which aggregates the values of stocks of 500 large companies traded on American stock exchanges and provides a credible view of the American stock market for investors.

There are a plethora of approaches toward recovering RND functions in the literature (see, for example, [5], for an extensive review). Parametric approaches typically specify a statistical model for the RND and the structural parameters are recovered by solving an optimization problem. For instance, a lognormal distribution was used in [6]; a mixture of lognormal distributions proposed by

[7] was considered in [2]; and a three-parameter Burr distribution was employed in [8], called the Burr family, which covers a broad range of shapes, including distributions similar to gamma, lognormal, and J-shaped beta. Another commonly used probability distribution in the literature of derivative pricing is the generalized hyperbolic distribution that contains variance gamma, normal inverse Gaussian (NIG), and t distributions as special cases (see, for instance, [9, 10]).

Nonparametric procedures, by contrast, are free from distributional assumptions on the underlying asset and thus achieve more flexibility than parametric methods. For example, in [11], cubic spline functions to model the unknown RND were used. An estimated density is numerically obtained by solving a quadratic programming problem with a convex objective function and nonnegativity constraints. The authors deliberately chose more knots than option strikes for higher flexibility. Lee (2014) [12] approximated the risk-neutral cumulative distribution function using quartic B-splines with power tails and the minimum number of knots that meet zero bid-ask spread. Their estimation was based on out-of-the-money option prices.

In this article, we propose a simpler but more powerful nonparametric solution using piecewise constant (PC) functions to estimate the RND. It is easy to implement since the estimating problem is formulated as a weighted least squared (WLS) procedure. It is more powerful since our method can recover the RND more effectively with all available option market prices without screening. Furthermore, our solution provides a practical way to explore profitable investment opportunities in financial markets by comparing the estimated prices and corresponding market prices.

The rest of this article is structured as follows. In *Materials and Methods*, we introduce the proposed nonparametric approach after reviewing cubic splines, quartic B-splines, and the NIG parametric approaches in the literature. In *Pricing European Options*, we run comprehensive cross-validation studies using the S&P 500 European option data to compare different methods and provide theoretical justifications on the consistency of our estimator for fair option prices. In *Pricing Variance Swap*, we apply the proposed nonparametric approach to price variance swaps, which is challenging in practice [13, 14]. We conclude in *Discussion*. The proofs and more formulae are collected in **the Supplementary Material.**

2 MATERIALS AND METHODS

In this section, we first provide a brief review of the cubic splines, quartic B-splines, and NIG approaches in the literature for recovering the RND. Then, we introduce the proposed PC nonparametric approach with least square (LS) and WLS procedures.

2.1 Nonnegative Cubic Spline Estimate for RND

Given the current trading date t and the expiration date T of European options, let $[K_1, K_q]$ be the range of strike prices of all available options traded in the market with the same underlying asset. Monteiro et al. (2008) [11] considered $s+1$ equally spaced knots for a cubic spline with $K_1 = x_1 < x_2 < x_3 < \cdots < x_s < x_{s+1} = K_q$. These knots are not necessarily a subset of the available strikes. Nevertheless, the closer these knots are to the strikes, the better their solution is. They also claimed that the number of knots should not be very much larger than the number of distinct strikes.

For the sake of nonnegativity of the estimated RNDs, in [11], the solution is much more complicated and computationally expensive than the usual cubic spline estimates. For comparison purposes, we keep only the constraints that ensure the nonnegativity of the density function on knots in their optimization procedure. By evaluating the difference between the estimated fair prices and the market prices of options, if our approach achieves higher accuracy than the cubic spline estimate with fewer constraints, then our approach is considered to be superior to that of [11].

When it comes to practical implementation [11], eliminating option prices is suggested that led to potential arbitrage opportunities according to the bid-ask interval, put-call parity, monotonicity, and strict convexity. They also generated "fake" call option prices using put-call parity to eliminate "artificial" arbitrage opportunities. Our comprehensive studies in *Pricing European Options* show that their screening and cleaning procedure may result in substantial information loss.

2.2 Quartic B-Spline Estimate

Lee (2014) [12] adopted a uniform quartic B-spline to estimate the risk-neutral cumulative distribution function (CDF). They used power tails to extrapolate outside the strike price range. They suggested using only the out-of-the-money (hereafter, OTM) options to estimate the CDF, including OTM call options whose strikes are higher than the underlying asset price and OTM put options whose strikes are lower than the underlying asset price. OTM options are typically cheaper than in-the-money (ITM) options and are considered to be more liquid as well. Nevertheless, our case studies in *Pricing European Options* show that ITM options may help recover the RND as well.

Due to fewer parameters, the quartic B-spline estimate is computationally more efficient than the nonnegative cubic spline approach. Lee (2014) [12] chose the number of knots needed as the minimum number that satisfies zero bid-ask pricing spread. They also suggested eliminating options that violate monotonicity and strict convexity constraints.

2.3 NIG Parametric Approach

For comparison purposes, we choose one parametric approach for approximating the RND, as suggested by [9, 15]. It is based on the NIG distribution, which belongs to the generalized hyperbolic class and can be characterized by its first four moments, i.e., mean, variance, skewness, and kurtosis. According to [16], these four moments can be estimated by the OTM European call and put options. One major issue with NIG density estimate is that, as shown in [10], the feasibility of NIG approach drops down as the time-to-maturity increases since more estimated skewness and kurtosis pairs fall outside the feasible domain of the NIG distribution.

2.4 The Proposed Piecewise Constant Nonparametric Approach

The PC approach that we propose in this article is nonparametric by nature. It is simpler but more efficient. Let S_t and S_T stand for the current price of equity on day t and the future price on day T. To estimate the RND function $f_{\mathbb{Q}}$ of $\log(S_T)$ conditional on the information up to day t, we propose to use a PC function, or a step function, to approximate $f_{\mathbb{Q}}$, with all distinct strike prices as knots. The constants in the step function are estimated by solving an optimization problem subject to certain constraints. By forcing the constants to be nonnegative, the nonnegativity of the estimated RND is guaranteed.

To be precise, suppose that we have a collection of market prices of European put and call options that are traded on date t and expire on date T. Let $\{K_1, K_2, \ldots, K_q\}$ represent the distinct strikes in ascending order and $\mathcal{C}$ be the collection of indices for call options and $\mathcal{P}$ for put options. Then, $\mathcal{C} \cup \mathcal{P} = \{1, 2, \ldots, q\}$. Let $m = |\mathcal{C}|$ and $n = |\mathcal{P}|$ be the numbers of calls and puts, respectively. Then $m + n \geq q$.

Given a RND $f_{\mathbb{Q}}$, the fair prices of put option and call option with strike K_i are

$$P_i = \mathbb{E}_t^{\mathbb{Q}} e^{-R_{tT}} (K_i - S_T)_+ = e^{-R_{tT}} \int_{-\infty}^{\log K_i} (K_i - e^y) f_{\mathbb{Q}}(y) dy,$$
$$C_i = \mathbb{E}_t^{\mathbb{Q}} e^{-R_{tT}} (S_T - K_i)_+ = e^{-R_{tT}} \int_{\log K_i}^{\infty} (e^y - K_i) f_{\mathbb{Q}}(y) dy,$$

respectively, where R_{tT} stands for the cumulative risk-free interest rate from t to T; that is, \$1 on day t ends for sure with $e^{R_{tT}}$ dollars on day T. We denote by r_t the risk-free interest rate over the period $[t, t+1]$, which is obtained from the risk-free zero-coupon bond, and clearly $R_{tT} = \sum_{j=t}^{T-1} r_j$.

To account for the RND outside the range $[K_1, K_q]$, we add $K_0 = K_1/c_K$ and $K_{q+1} = c_K K_q$, where $c_K > 1$ is a predetermined constant that can be chosen by means of cross-validation or prior knowledge (see details in *Pricing European Options*). We then use a PC function f_Δ to approximate $f_{\mathbb{Q}}$; that is,

$$f_\Delta(y) = a_l, \quad \text{for} \quad \log K_{l-1} < y \leq \log K_l, \quad l = 1, 2, \ldots, q+1, \tag{1}$$

and zero elsewhere. Here, $\Delta = \{\log K_1, \ldots, \log K_q\}$ stands for the collection of distinct strikes in log scale and $\{a_l, l = 1, \ldots, q+1\}$ are nonnegative constants satisfying

$$\sum_{l=1}^{q+1} a_l \log \frac{K_l}{K_{l-1}} = 1 \tag{2}$$

due to the condition $\int_{-\infty}^{+\infty} f_\Delta(y) dy = 1$.

Given the approximate RND f_Δ, the estimated put and call prices with strike K_i are

$$\widehat{P}_i = e^{-R_{tT}} \int_{-\infty}^{\log K_i} (K_i - e^y) f_\Delta(y) dy, \tag{3}$$

$$\widehat{C}_i = e^{-R_{tT}} \int_{\log K_i}^{\infty} (e^y - K_i) f_\Delta(y) dy, \tag{4}$$

respectively, which are essentially linear functions of $a_1, \ldots, a_q$.

Proposition 2.1. Given $a_l \geq 0, l = 1, \ldots, q+1$ satisfying **Eq. 2**, the estimated prices for put and call options with strike K_i satisfy

$$e^{R_{tT}} \widehat{P}_i = a_1 X_{i,1}^{(P)} + \cdots + a_q X_{i,q}^{(P)} + X_{i,q+1}^{(P)}, \tag{5}$$

$$e^{R_{tT}} \widehat{C}_i = a_1 X_{i,1}^{(C)} + \cdots + a_q X_{i,q}^{(C)} + X_{i,q+1}^{(C)}, \tag{6}$$

where $X_{i,l}^{(P)} = X_{i,l}^{(p)} - \log(K_l/K_{l-1}) (\log c_K)^{-1} X_{i,q+1}^{(p)}$, $X_{i,l}^{(C)} = X_{i,l}^{(c)} - \log(K_l/K_{l-1}) (\log c_K)^{-1} X_{i,q+1}^{(c)}$, $l = 1, 2, \ldots, q$; $X_{i,q+1}^{(P)} = X_{i,q+1}^{(p)} (\log c_K)^{-1}$, $X_{i,q+1}^{(C)} = X_{i,q+1}^{(c)} (\log c_K)^{-1}$; *and* $X_{i,l}^{(p)} = [K_i \log(K_l/K_{l-1}) - (K_l/K_{l-1})] \cdot 1(K_i \geq K_l)$, $X_{i,l}^{(c)} = [(K_l - K_{l-1}) - K_i \log(K_l - K_{l-1})] \cdot 1(K_i < K_l)$, $l = 1, 2, \ldots, q+1$.

The proof of Proposition 2.1 is relegated to **Supplementary Material A**.

The unknown parameters $a_1, \ldots, a_{q+1}$ are estimated by minimizing the following LS objective function:

$$L(a_1, \ldots, a_{q+1}) = \frac{1}{m+n} \left[\sum_{i \in \mathcal{C}} \left(\widehat{C}_i - \tilde{C}_i \right)^2 + \sum_{i \in \mathcal{P}} \left(\widehat{P}_i - \tilde{P}_i \right)^2 \right] \tag{7}$$

subject to $a_l \geq 0, l = 1, 2, \ldots, q+1$, and **Eq. 2**, where $\tilde{C}_i$ and $\tilde{P}_i$ are market prices of call option and put option, respectively, with strike K_i . If there exists an RND $f_{\mathbb{Q}}$, we have $C_i = \tilde{C}_i, i \in \mathcal{C}$ and $P_i = \tilde{P}_i, i \in \mathcal{P}$. That is, market prices are fair if there is no arbitrage in the financial market.

From an investment point of view, because a more expensive option tends to be less liquid, an alternative approach to determining $a_1, \ldots, a_{q+1}$ is to minimize a WLS objective function:

$$W(a_1, \ldots, a_{q+1}) = \frac{1}{m+n} \left[\sum_{i \in \mathcal{C}} \left(\frac{\widehat{C}_i - \tilde{C}_i}{\tilde{C}_i} \right)^2 + \sum_{i \in \mathcal{P}} \left(\frac{\widehat{P}_i - \tilde{P}_i}{\tilde{P}_i} \right)^2 \right]. \tag{8}$$

The WLS estimate is in favor of OTM options over ITM options in that OTM options are typically less expensive and more liquid.

3 PRICING EUROPEAN OPTIONS

In this section, we use the S&P 500 European options to evaluate the performances of various RND estimators.

3.1 S&P 500 European Option Data

We consider European calls and puts written on the S&P 500 indices from January 2, 1996, to August 31, 2015, in the United States [10, 14]. The expiration dates are the third Saturday of the delivery month. Following [14], we keep only the options with positive bid prices and positive volumes and with expiration date of more than seven days in our analysis. Similar to [10], we categorize options into seven groups with expiration in 7 ~ 14, 17 ~ 31, 81 ~ 94, 171 ~ 199, 337 ~ 393, 502 ~ 592, and 670 ~ 790 days, respectively, for the purpose of examining the effects of the length of maturity on pricing. The numbers of options and (t, T) pairs under consideration are presented in **Table 1**.

TABLE 1 | Numbers of calls, puts, and (t, T) pairs in different time-to-maturity categories (number of days to expiration).

#Day	7~14	17~31	81~94	171~199	337~393	502~592	670~790
#Call	72,535	136,019	34,764	17,367	13,465	7,985	5,869
#Put	112,862	205,863	53,648	27,906	18,982	14,535	10,104
#(t, T)	2,411	4,206	2,548	2,306	2,747	2,536	1,739

3.2 Comprehensive Comparisons With Existing Methods

We use the S&P 500 European options to evaluate the performance of the following methods: the parametric NIG estimate, the quartic B-spline (B-spline) estimate, the nonnegative cubic spline estimates with either LS criterion (cubic + LS) or WLS criterion (cubic + WLS), and the proposed PC estimate with either LS or WLS objective function using OTM options only (PC + LS + OTM or PC + WLS + OTM) or using all available options (PC + LS + ALL or PC + WLS + ALL). All the comparisons are made based on their ability to recover option market prices.

For each of the seven time-to-maturity categories listed in **Table 1**, we randomly selected 200 pairs of (t, T). For each pair, the market prices of calls and puts are collected. The aforementioned approaches are applied to estimate the RND of the underlying asset at time T. We then use the estimated RND to obtain $\widehat{C}_i$ and $\widehat{P}_i$. The discrepancy between the market prices and the estimated prices is assessed by means of the absolute error L_a and the relative error L_r defined as follows:

$$L_a^2 = \frac{1}{|\mathcal{C}_t| + |\mathcal{P}_t|}\left[\sum_{i\in\mathcal{C}_t}\left(\widehat{C}_i - \tilde{C}_i\right)^2 + \sum_{i\in\mathcal{P}_t}\left(\widehat{P}_i - \tilde{P}_i\right)^2\right],$$
$$L_r^2 = \frac{1}{|\mathcal{C}_t| + |\mathcal{P}_t|}\left[\sum_{i\in\mathcal{C}_t}\left(\widehat{C}_i/\tilde{C}_i - 1\right)^2 + \sum_{i\in\mathcal{P}_t}\left(\widehat{P}_i/\tilde{P}_i - 1\right)^2\right],$$

where $\mathcal{C}_t$ (or $\mathcal{P}_t$) refers to the collection of indices of call (or put) options used for testing purposes. In **Table 2** and **Table 3**, we choose $\mathcal{C}_t$ and $\mathcal{P}_t$ to be either all available OTM options or ITM options. We report the average L_a and L_r over the 200 randomly selected pairs of (t, T) for each estimation approach. The columns labeled '200' show the actual number of pairs that yield a valid RND estimate. The higher the count, the more effective the method. As explained in Section 2.3, the NIG approach is quite picky in selecting calls and puts. For B-spline and cubic methods, following the same filtering procedures as in [11, 12], respectively, we observe that fewer options become available as the time-to-maturity increases, which results in substantial information loss. On the contrary, our PC methods with LS or WLS are feasible for almost all cases, especially when using both ITM and OTM options.

In terms of the absolute error L_a and the relative error L_r computed for different combinations of time-to-maturities and RND estimates, our PC estimates are more stable and accurate than the other three approaches. As illustrated in **Tables 2, 3**, the proposed PC methods always yield the lowest L_a or L_r, regardless of the type of options used. In order for a cross-sectional comparison to be conducted among all the approaches, only OTM options are considered when using the proposed PC approach to price options (i.e., PC + LS + OTM or PC + WLS + OTM). In practice, however, we would recommend using all available option prices, including both ITM and OTM options. In particular, if the goal is to obtain the most precise price, we recommend 'PC + LS + ALL' in that it controls absolute error L_a the best; if one seeks a higher return on investment, we would recommend 'PC + WLS + ALL' instead, which controls relative error L_r the best.

3.3 Consistency of PC Estimates for Fair Prices

Given distinct strike prices $K_1 < K_2 < \cdots < K_q$, the associated market prices of calls and puts, $\{\tilde{C}_i, i \in \mathcal{C}\}$ and $\{\tilde{P}_i, i \in \mathcal{P}\}$, respectively, traded on date t with expiration date T satisfy

TABLE 2 | Comprehensive comparison of different RND estimates: part I.

Time-to-maturity		7~14			17~31			81~94			171~199		
Method	Test	L_a	L_r	200	L_a	L_r	200	L_a	L_r	200	L_a	L_r	200
NIG	ITM	1.823	0.058	145	2.293	0.057	110	5.902	0.095	91	14.344	0.158	143
	OTM	0.772	0.569	145	1.669	0.533	110	5.404	0.769	92	10.445	0.771	143
B-spline	ITM	27.031	0.107	133	30.404	0.140	156	33.019	0.124	93	23.086	0.144	30
	OTM	1.638	15.102	133	9.981	64.444	156	5.037	7.153	93	11.783	12.655	30
Cubic + LS	ITM	3.645	0.218	102	1.055	0.028	76	2.254	0.041	77	69,861.9	734.427	69
	OTM	3.105	4.452	102	0.387	0.600	76	1.094	1.276	77	224,532.5	15,259.638	69
Cubic + WLS	ITM	4.696	0.236	102	1.286	0.034	76	2.977	0.049	77	66,506.6	699.154	69
	OTM	3.480	5.032	102	0.446	0.656	76	1.297	1.084	77	214,119.1	14,806.153	69
PC + LS + ALL	ITM	0.138	0.005	200	0.150	0.004	200	0.269	0.004	200	0.430	0.004	200
	OTM	0.083	0.157	200	0.097	0.114	200	0.162	0.077	200	0.420	0.056	200
PC + WLS + ALL	ITM	0.219	0.007	200	0.231	0.005	200	0.462	0.006	200	1.628	0.008	200
	OTM	0.077	0.074	200	0.090	0.064	200	0.166	0.034	200	0.370	0.028	200
PC + LS + OTM	ITM	0.679	0.023	200	0.646	0.015	200	6.570	0.098	198	25.042	0.171	198
	OTM	0.053	0.098	200	0.074	0.086	200	0.153	0.043	198	0.275	0.036	198
PC + WLS + OTM	ITM	0.913	0.029	200	0.803	0.019	200	9.073	0.114	198	25.135	0.172	198
	OTM	0.121	0.077	200	0.121	0.065	200	0.308	0.034	198	0.364	0.025	198

TABLE 3 | Comprehensive comparison of different RND estimates: part II.

Time-to-maturity		337~393			502~592			670~790		
Method	**Test**	L_a	L_r	200	L_a	L_r	200	L_a	L_r	200
NIG	ITM	23.238	0.165	51	37.368	0.216	27	49.308	0.180	29
	OTM	17.781	0.790	53	28.575	1.236	27	33.174	5.329	29
B-spline	ITM	146.941	0.255	4	NA	NA	0	NA	NA	0
	OTM	146.941	0.255	4	NA	NA	0	NA	NA	0
Cubic + LS	ITM	251,615.4	876.7	68	248,235.4	778.5	75	47,327.8	95.317	54
	OTM	110,639.6	1,553.4	68	303,111.7	24,539.1	75	24,203.2	2,351.626	54
Cubic + WLS	ITM	250,487.1	872.8	68	406,119.5	1,259.7	75	47,364.3	95.391	54
	OTM	110,189.3	1,547.4	68	517,077.6	35,028.3	75	24,205.7	2,353.728	54
PC + LS + ALL	ITM	4.907	0.033	200	7.406	0.089	200	7.501	0.062	200
	OTM	1.323	0.066	200	3.100	0.154	200	5.484	0.098	200
PC + WLS + ALL	ITM	7.148	0.035	200	7.778	0.070	200	6.556	0.051	200
	OTM	2.256	0.028	200	4.382	0.044	200	6.268	0.048	200
PC + LS + OTM	ITM	79.914	0.320	192	92.636	0.439	197	86.013	0.404	194
	OTM	1.150	0.032	192	1.401	0.050	197	1.153	0.044	194
PC + WLS + OTM	ITM	79.688	0.318	192	93.565	0.438	197	85.833	0.403	194
	OTM	1.461	0.021	192	2.1977	0.021	197	1.479	0.018	194

$$\tilde{C}_i = e^{-R_t T}\int_{\log K_i}^{\infty}(e^y - K_i) f_{\mathbb{Q}}(y)dy,$$
$$\tilde{P}_i = e^{-R_t T}\int_{-\infty}^{\log K_i}(K_i - e^y) f_{\mathbb{Q}}(y)dy, \tag{9}$$

provided that RND $f_{\mathbb{Q}}$ of $\log S_T$ exists. That is, the market prices $(\tilde{C}_i, \tilde{P}_i)$ agree with the fair prices (C_i, P_i) .

According to **Eq. 1**, the proposed PC approach provides the following approximation:

$$f_\Delta(x) = \sum_{l=1}^{q+1} a_l 1_{(\log K_{l-1}, \log K_l]}(x) \tag{10}$$

to the RND $f_{\mathbb{Q}}$, where $(a_1, \ldots, a_{q+1})$ minimizes the absolute error $L(a_1, \ldots, a_{q+1})$ or the relative error $W(a_1, \ldots, a_{q+1})$. The estimated fair prices $(\widehat{P}_i, \widehat{C}_i)$ calibrated using f_Δ are determined by **Eqs 3, 4**.

Because $f_{\mathbb{Q}}$ is often not unique in practice, instead of measuring the distance between f_Δ and $f_{\mathbb{Q}}$, we would like to ask whether the prices obtained using f_Δ could recover the market prices well. The extensive numerical studies reported in **Tables 2, 3** corroborate this claim. This is further justified by the following theorem.

Theorem 3.1. *Suppose there exists a continuous RND $f_{\mathbb{Q}}$ of $\log S_T$, satisfying $\int_0^\infty e^x f_{\mathbb{Q}}(x)dx < \infty$. Let $\Delta = \{\log K_1, \ldots, \log K_q\}$ be the collection of distinct strike prices in log scale with both call and put option market prices available. Then as $K_1 \to 0$, $K_q \to \infty$, $q \to \infty$, and $|\Delta| := \max_{1 \le i < q} \log(K_{i+1}/K_i) \to 0$, we have*

$$\frac{1}{2q}\left[\sum_{i=1}^{q}\left(\widehat{C}_i - \tilde{C}_i\right)^2 + \sum_{i=1}^{q}\left(\widehat{P}_i - \tilde{P}_i\right)^2\right] \to 0.$$

Remark 1. Since $\tilde{C}_i = e^{-R_t T}\int_{\log K_i}^{\infty} e^y f_{\mathbb{Q}}(y)dy - K_i e^{-R_t T}\int_{\log K_i}^{\infty} f_{\mathbb{Q}}(y)dy$, then the condition $\int_0^\infty e^x f_{\mathbb{Q}}(x)\,dx < \infty$ in Theorem 3.1 is necessary and sufficient for $\tilde{C}_i < \infty$.

Remark 2. The proof for Theorem 3.1 is relegated to **Supplementary Material B**. It shows the existence of $(a_1, \ldots, a_{q+1})$ such that $\max_{1 \le i \le q}\left|\widehat{C}_i - \tilde{C}_i\right| < \epsilon$ and $\max_{1 \le i \le q}\left|\widehat{P}_i - \tilde{P}_i\right| < \epsilon$ for any given $\epsilon > 0$ when $K_1, |\Delta|$ are sufficiently small and K_q, q are sufficiently large. In other words, $\left|\widehat{C}_i - \tilde{C}_i\right|, \left|\widehat{P}_i - \tilde{P}_i\right|, i = 1, \ldots, q$, can be uniformly small.

3.4 Detecting Profitable Opportunities

Theorem 3.1 provides analytical foundations for the consistency of the proposed PC method under the assumption of the existence of a continuous RND. Nevertheless, the PC method is still applicable even when there is an arbitrage opportunity in the market. In this case, a significant difference between the market price and its estimated fair price would be expected.

With a given set of market option prices, our nonparametric method can recover a fair option price for any strike price. From an investment point of view, we are able to detect options on the markets that are under- or overpriced. It may not be adequate to claim arbitrage opportunities due to the lack of guarantee to earn and since there is a mature market system designed to catch such kind of difference among the option prices. Nevertheless, we can still report profitable investment opportunities for investors.

In **Figure 1**, we provide an illustrative example using $m + n = 95$ available market prices of options traded on April 14, 2014 with expiration September 5, 2014. For each of the 95 options, we obtain its fair price based on our PC + LS method using the market prices of the rest $m + n - 1 = 94$ options. Then, we compare the market price and the leave-one-out fair price, known as leave-one-out cross-validation. **Figure 1A** depicts $m + n = 95$ market prices in dots and leave-one-out fair prices in solid line against the corresponding strike prices. It seems that they match each other very well.

To have a closer look at the difference between market price and fair price, we plot the relative difference, that is (market price - fair price)/fair price, against strike price in **Figure 1B**. The sign of the relative difference tells us whether the option is under- or overpriced. In addition, in order to check if the difference between a market price and its fair price is statistically

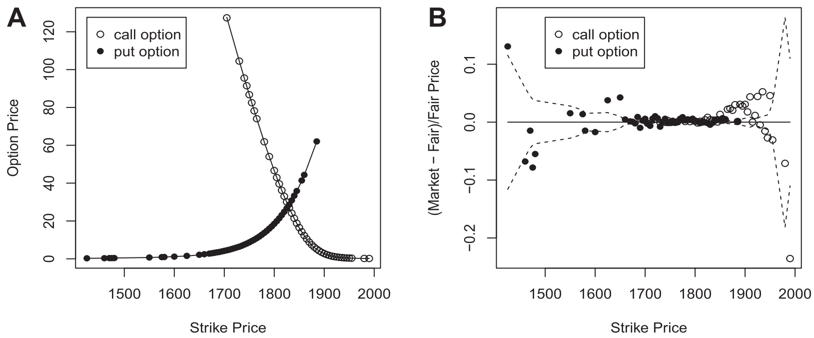

FIGURE 1 | Leave-one-out cross-validation for options traded on April 14, 2014 with an expiration date of September 5, 2014 (round dot: market price; solid line: fair price based on PC; dash line: 95% confidence interval based on bootstrap).

significant, we bootstrap the rest of the market prices 50 times to obtain a 95% confidence interval of the fair price. The dash lines in **Figure 1B** show the upper and lower ends of the bootstrap confidence intervals. When a market price falls outside its bootstrap confidence interval, we may report to investors that the corresponding option is significantly under-/overpriced compared with the market prices of the other options.

4 PRICING VARIANCE SWAP

With an estimated RND, one can calculate the fair price of any derivative whose payoff is a function of S_T. In this section, we apply the proposed method to price variance swaps. Our study shows that our fair prices match the market prices of long-term variance swaps reasonably well.

A variance swap is a financial product that allows investors to trade realized variance against current implied variance of log returns. More specifically, let S_t stand for the closing price of the underlying asset on day t, $t = 0, 1, \ldots, T$, and let $R_t = \log(S_t/S_{t-1})$ represent the tth daily log return. The annualized realized variance over T trading days is defined as $\sigma^2_{\text{realized}} = \frac{A}{T}\sum_{t=1}^{T} R_t^2$, where A is the number of trading days per year, which on average is 252. The payoff of a variance swap is defined as

$$N_{\text{var}}\left(\sigma^2_{\text{realized}} - \sigma^2_{\text{strike}}\right),$$

where the variance notional N_{var} and variance strike σ^2_{strike} are specified before the sale of a variance swap contract.

Variance swaps provide investors with pure exposure to the variance of the underlying asset without directional risk. It is notably liquid across major equities, indices, and stock markets and is growing across other markets. Historical evidence indicates that selling variance is systematically profitable.

There are numerous methods in the literature of pricing variance swaps, both analytically and numerically (see [13] for an extensive review). Nevertheless, a pricing formula or procedure that relies on a certain stochastic process, for instance, Lévy process [14], MRG-Vasicek model [17], or Hawkes jump-diffusion model [18], may suffer from a lack of parsimony or might not fit the real data well due to the inappropriateness of model assumptions (see, for instance, [14]).

In this section, we propose a moment-based method in conjunction with our PC RND estimate to price variance swaps, which is free of model assumption.

Assuming the existence of a risk-neutral measure $\mathbb{Q}$, the fair price $VS_{t,T}$ of a variance swap on day t is the discounted expected payoff:

$$VS_{t,T} = e^{-R_{tT}} N_{\text{var}}\left[\mathbb{E}_t^{\mathbb{Q}}\left(\frac{A}{T}\sum_{i=1}^{T} R_i^2\right) - \sigma^2_{\text{strike}}\right], \tag{11}$$

where R_{tT} is the cumulative risk-free interest rate from t to T defined in *The Proposed Piecewise Constant Nonparametric Approach*. To proceed, we further assume the following.

Assumption 1. The increments of the process $\log S_t$ are independent; that is, $\log(S_{t+1}/S_t)$ is independent of $S_0, \ldots, S_t$, $t = 0, \ldots, T-1$.

Consequently, the fair price of a variance swap can be represented by a sequence of the risk-neutral moments of the underlying asset.

Proposition 4.1. Assuming the existence of a risk-neutral measure $\mathbb{Q}$ and that Assumption 1 is fulfilled, the fair price of variance swap is

$$VS_{t,T} = e^{-R_{tT}} N_{\text{var}}\left\{\frac{A}{T}\sum_{i=1}^{t} R_i^2 + \frac{A}{T}\mathbb{E}_t^{\mathbb{Q}}(\log S_T)^2 - \frac{A}{T}(\log S_t)^2 - \frac{2A}{T}\sum_{i=t+1}^{T}\left[\mathbb{E}_t^{\mathbb{Q}}\log S_{i-1}\mathbb{E}_t^{\mathbb{Q}}\log S_i - \left(\mathbb{E}_t^{\mathbb{Q}}\log S_{i-1}\right)^2\right] - \sigma^2_{\text{strike}}\right\}. \tag{12}$$

The proof of Proposition 4.1 is relegated to **Supplementary Material C**.

4.1 Moments Calculation

In view of **Eq. 12**, pricing variance swaps requires estimating the first and second moments of $\log S_i$ under the risk-neutral measure. One option is to use a moment-based method described by [16] and further extended by [19]. In this section, we employ an alternative way of calculating the moments, which makes use of the proposed nonparametric approach.

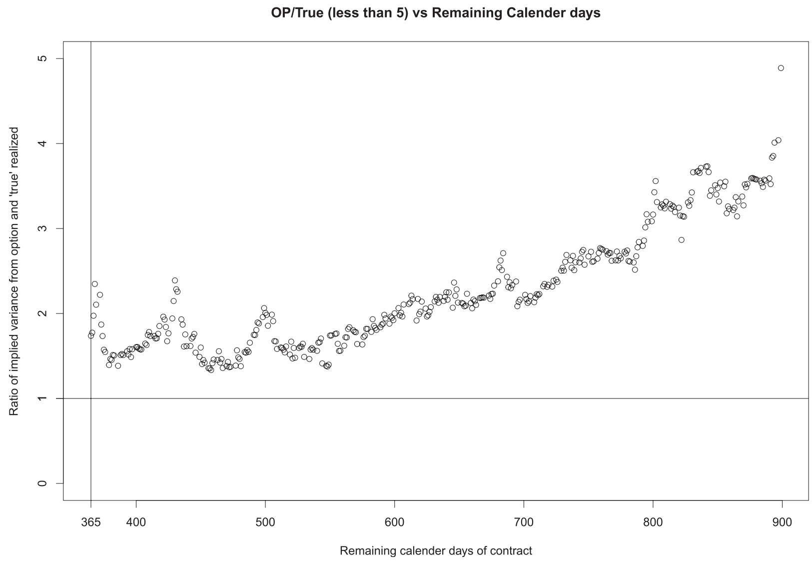

FIGURE 2 | Ratio of OP/True variance swap prices vs. days to expiration.

Recall that the step function f_Δ defined in **Eq. 1** or **Eq. 10** provides an approximation to the RND $f_{\mathbb{Q}}$ of $\log S_T$. We use all the available market prices of options to estimate f_Δ; then, the moments calculated from f_Δ serve as the estimates of required moments. Since f_Δ is "PC" which stands for piecewise constant. it can be verified that the first and second moments of $\log(S_T)$ are given by

$$\mathbb{E}_t^{\mathbb{Q}}\log(S_T) = \sum_{l=1}^{q+1} \frac{a_l}{2}\left[(\log K_l)^2 - (\log K_{l-1})^2\right], \tag{13}$$

$$\mathbb{E}_t^{\mathbb{Q}}[\log(S_T)]^2 = \sum_{l=1}^{q+1} \frac{a_l}{3}\left[(\log K_l)^3 - (\log K_{l-1})^3\right]. \tag{14}$$

Note that there are no market prices available for options that expire on a day that is other than the third Saturday of the delivery month. We would have to interpolate the mean and standard deviation of $\log(S_i)$ for $t < i < T$, and this is achieved via linear interpolation in this paper. Detailed procedures are described in **Supplementary Material D**.

4.2 Replicating by Variance Futures

In order to evaluate the fair price of a variance swap, we replicate variance swap using available market prices of variance futures. Variance future is a financial contract that is traded over the counter. As stated in [20], variance swap and variance future are essentially the same since they both trade the difference of variance and one can replicate a variance swap by the corresponding variance future. As a matter of fact, if variance future and variance swap share the same expiration date, then at the start point of the observation period, there is no difference between trading a variance future and trading a variance swap with $50 variance notional. The formula for the fair price of a variance swap contract induced from variance future is given by

$$VS_{t,T} = e^{-R_t T} N_{\text{var}} \left\{ \frac{A}{T}\left[\sum_{i=1}^{M-1} R_i^2 + IUG \times \frac{N_e - M + 1}{A} \times \frac{1}{100^2}\right] - \sigma_{\text{strike}}^2 \right\},$$

where M is the number of observed days to date, N_e is the expected number of trading days in the observation period, and IUG is the square of market implied volatility given by

$$IUG = \sum_{i=M}^{N_e} R_i^2 \times \frac{A}{N_e - M + 1} \times 100^2.$$

4.3 Variance Future Data

Variance future data were downloaded from the Chicago Board Options Exchange (CBOE) website (http://cfe.cboe.com/). Variance future products with 12-month (with futures symbol VA) or 3-month (with futures symbol VT) expirations are traded on the CBOE Futures Exchange. We use VA in the subsequent analysis. The continuously compounded zero-coupon interest rates cover dates from January 2, 1996 to August 31, 2015. For variance futures, the trading dates are from December 10, 2012, to August 31, 2015, with start dates from December 21, 2010, to July 30, 2015, and expiration dates from January 18, 2013, to January 1, 2016. We use variance futures to replicate variance swaps, so the time spans of variance swaps are in line with those of variance futures.

4.4 Results

In order to assess the accuracy of our estimated fair prices of a variance swap, we compare three relevant quantities:

1. OP: Fair price of a variance swap based on our moment-based nonparametric approach, using option market prices till day t
2. VF: Induced market price of a variance swap from CBOE traded variance future till day t

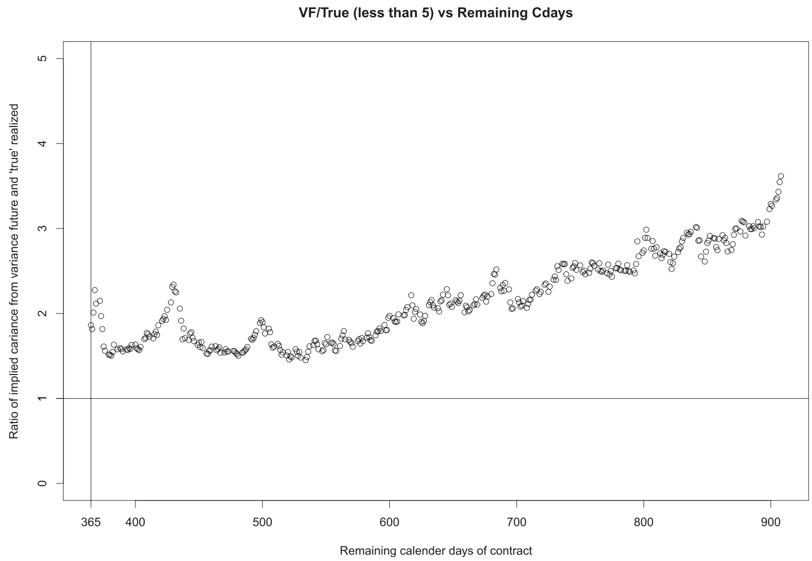

FIGURE 3 | Ratio of VF/True variance swap prices vs. days to expiration.

3. True: Realized price of a variance swap at expiration day T with known $S_0, S_1, \ldots, S_T$

We present three ratios, OP/True, VF/True, and OP/VF, in **Figures 2–4**, respectively, against the remaining calendar days of variance swaps. Compared with 'True' prices based on realized underlying asset prices, **Figure 2** and **Figure 3** suggest that OP and VF have a similar increasing pattern along with the remaining calendar days. This is in part due to the uncertainty in the estimate of the variance, which increases with the number of days to expiration. On the other hand, **Figure 4** and **Figure 5** show that the fair price OP based on our proposed method matches the market price VF pretty well on variance swaps with expiration between 365 and 800 days. For variance swaps expiring in less than 365 days (not shown here), OP and VF do not match well. This is plausibly attributed to the fact that long-term options are more reasonable and stable, which are less likely to be affected by external factors or noises. For variance swaps longer than 800 days, the relatively low VF might indicate underpriced variance futures.

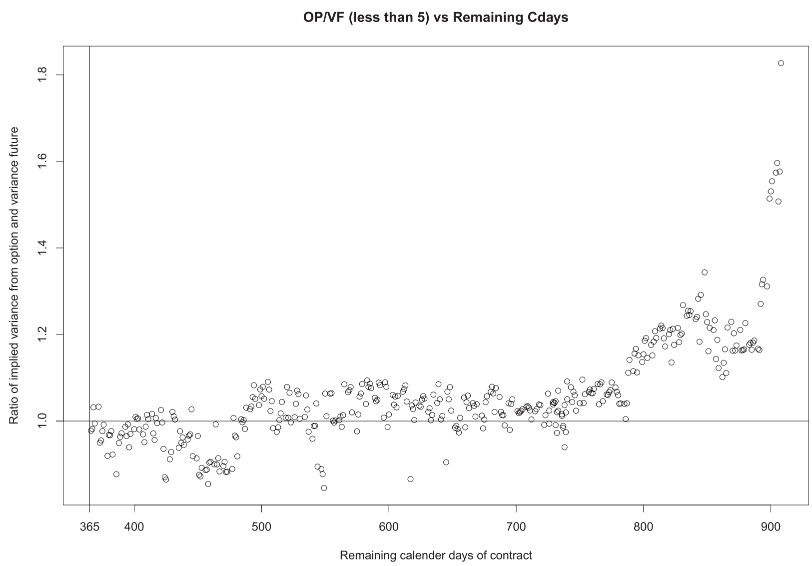

FIGURE 4 | Ratio of OP/VF variance swap prices vs. days to expiration.

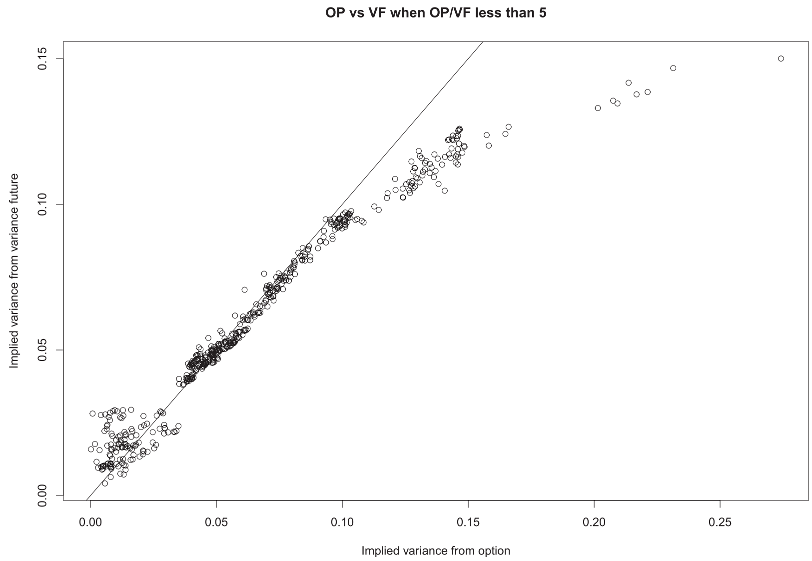

FIGURE 5 | OP vs. VF variance swap prices.

5 DISCUSSION

In this article, we propose a new nonparametric approach for estimating the RND. It is data-driven and is not built on any model assumption about the data generating process of underlying asset prices. It only assumes the existence of an RND and the independence of increments of log return for pricing variance swaps. That is why it can capture the market price very well.

In contrast with other nonparametric methods, such as cubic spline and B-spline, our method is much simpler but fits the real data better. We choose only distinct strikes as knots and assume constant values between knots to avoid overfitting. By sacrificing the continuity of estimated RND, the nonnegativity of a density function is readily satisfied.

On the other hand, the proposed approach utilizes market prices of all options, not just OTM options. In our opinion, ITM options, despite not being as liquid as OTM options, still contain market information and should be incorporated when estimating a RND. Our comprehensive analysis shows that it recovers OTM option prices better by including ITM option prices.

One of the potential applications of our work is to price OTC securities. The comprehensive comparisons of different methods using S&P 500 European option data in *Pricing European Options* show the outperformance of our estimates across various time-to-maturity periods. With our estimated RNDs, financial engineers may be able to develop more sophisticated financial products to fit the needs of their customers better. A fair price of the new financial product will not only facilitate the traders but also reduce their financial risk.

Another potential application is to build up profitable portfolios for investment purposes. In *Detecting Profitable Opportunities*, we show how to use leave-one-out and bootstrap techniques to identify under- or overpriced options from fair-priced options. Real-time trading algorithms may be developed based on our estimates toward catching up profitable opportunities by buying underpriced options and shorting overpriced options.

Pricing variance swaps is a difficult job when dealing with real data. We display in **Figures 2–5** only the cases where the ratio OP/True is less than 5. There are cases where OP and VF disagree significantly. Overall, our OP prices work better for variance futures that expire in the last four months of 2015, which are also the last four months available in our dataset.

In the literature, RND is naturally connected to the concept of statistic discount factor or pricing kernel, which can be described as the ratio of the RND and the physical density [21]. The proposed nonparametric estimate of RND could be applied to estimate pricing kernel as well.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: http://cfe.cboe.com/, Chicago Board Options Exchange (CBOE) website.

AUTHOR CONTRIBUTIONS

LJ and SZ processed the experimental data and performed the analysis. LJ drafted the manuscript. JY, LJ, and KL developed the theory and performed the analytical calculations. FW and KL helped supervise the project. JY, FW, and LJ conceived the original idea. JY supervised the project. All authors discussed the results and contributed to the final manuscript.

ACKNOWLEDGMENTS

We thank Liming Feng from the University of Illinois at Urbana-Champaign and Yuhang Liang from Northwestern University for their enormous help during data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fams.2020.611878/full#supplementary-material.

REFERENCES

1. Malz AM. *A simple and reliable way to compute option-based risk-neutral distributions*. In: *Staff report* New York, NY: Federal Reserve Bank of New York. (2014) p 677.
2. Melick W, Thomas C. Recovering an asset's implied pdf from option prices: an application to crude oil during the gulf crisis. *J Financ Quant Anal* (1997) 32: 91–115. doi:10.2307/2331318
3. Taboga M. Option-implied probability distributions: how reliable? How jagged? *Int Rev Econ Finance* (2016) 45:453–69. doi:10.1016/j.iref.2016.07.013
4. Breeden D, Litzenberger R. Prices of state-contingent claims implicit in option prices. *J Bus* (1978) 51:621–51.
5. Bliss R, Panigirtzoglou N. Testing the stability of implied probability density functions. *J Bank Finance* (2002) 26:381–422. doi:10.2139/ssrn.209650
6. Jarrow R, Rudd A. Approximate option valuation for arbitrary stochastic process. *J Financ Econ* (1982) 10:347–69. doi:10.1016/0304-405X(82)90007-1
7. Ritchey R. Call option valuation for discrete normal mixtures. *J Financ Res* (1990) 13:285–96. doi:10.1111/j.1475-6803.1990.tb00633.x
8. Sherrick B, Irwin S, Forster D. Option-based evidence of the nonstationarity of expected SP 500 futures price distributions. *J Futures Mark* (1992) 12:275–90. doi:10.1002/fut.3990120304
9. Eriksson A, Ghysels E, Wang F. The normal inverse Gaussian distribution and the pricing of derivatives. *J Deriv* (2009) 16:23–37. doi:10.3905/JOD.2009.16.3.023
10. Ghysels E, Wang F. Moment-implied densities: properties and applications. *J Bus Econ Stat* (2014) 32:88–111. doi:10.1080/07350015.2013.847842
11. Monteiro A, Tutuncu R, Vicente L. Recovering risk-neutral probability density functions from options prices using cubic splines. *Eur J Oper Res* (2008) 187: 525–42. doi:10.1016/j.ejor.2007.02.041
12. Lee S. Estimation of risk-neutral measures using quartic b-spline cumulative distribution functions with power tails. *Quant Finance* (2014) 14:1857–79. doi:10.1080/14697688.2012.742202
13. Zhu SP, Lian GH. A closed-form exact solution for pricing variance swaps with stochastic volatility. *Math Finance* (2010) 21:233–56. doi:10.1111/j.1467-9965.2010.00436.x
14. Carr P, Lee R, Wu L. Variance swaps on time-changed Lévy processes. *Finance Stochast* (2012) 16:335–55. doi:10.1007/s00780-011-0157-9
15. Eriksson A, Ghysels E, Forsberg L. *Approximating the probability distribution of functions of random variables: a new approach*. Montreal, Canada:Cirano (2004) p 503.
16. Bakshi G, Kapadia N, Madan D. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Rev Financ Stud* (2003) 16: 101–43. doi:10.2139/srn.282451
17. Han Y, Zhao L. A closed-form pricing formula for variance swaps under MRG-Vasicek model. *Comput Appl Math* (2019) 38:142. doi:10.1007/s40314-019-0905-6
18. Liu W, Zhu SP. Pricing variance swaps under the Hawkes jump-diffusion process. *J Futures Mark* (2019) 39:635–55. doi:10.1002/fut.21997
19. Rompolis LS, Tzavalis E. Retrieving risk neutral moments and expected quadratic variation from option prices. *Rev Quant Finance Account* (2017) 48:955–1002. doi:10.2139/ssrn.936670
20. Biscamp L, Weithers T. *Variance swaps and cboe s and p 500 variance futures*. Chicago Trading Company, LLC. (2007) p 25.
21. Beare BK, Schmidt LD. An empirical test of pricing kernel monotonicity. *J Appl Econom* (2016) 31:338–56. doi:10.1002/jae.2422

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Material
for
A New Nonparametric Estimate of the Risk-Neutral Density with Applications to Variance Swaps

A. PROOF OF PROPOSITION 2.1

We rewrite the call and put option prices in **Eqs 3, 4** in terms of $a_1, a_2, \ldots, a_q, a_{q+1}$ as follows

$$\begin{aligned} e^{R_{tT}}\hat{P}_i &= \int_{-\infty}^{\log K_i} (K_i - e^y) f_\Delta(y)dy \\ &= \sum_{l=1}^{q+1} \int_{\log K_{l-1}}^{\log K_l} (K_i - e^y) a_l dy \cdot \mathbb{1}(K_i \geq K_l) \\ &= \sum_{l=1}^{q+1} a_l[(K_i \log \frac{K_l}{K_{l-1}}) - (K_l - K_{l-1})] \cdot \mathbb{1}(K_i \geq K_l),\ i \in \mathcal{P} \end{aligned} \tag{S1}$$

$$\begin{aligned} e^{R_{tT}}\hat{C}_i &= \int_{\log K_i}^{\infty} (e^y - K_i) f_\Delta(y)dy \\ &= \sum_{l=1}^{q+1} \int_{\log K_{l-1}}^{\log K_l} (e^y - K_i) a_l dy \cdot \mathbb{1}(K_i \leq K_{l-1}) \\ &= \sum_{l=1}^{q+1} a_l[(K_l - K_{l-1}) - K_i \log \frac{K_l}{K_{l-1}}] \cdot \mathbb{1}(K_i < K_l),\ i \in \mathcal{C} \end{aligned} \tag{S2}$$

Let $X^{(p)}_{i,l} = [K_i \log(K_l/K_{l-1}) - (K_l - K_{l-1})] \cdot \mathbb{1}(K_i \geq K_l)$, $l = 1, 2, \ldots, q+1$ be an entry of the design matrix for put options; and $X^{(c)}_{i,l} = [(K_l - K_{l-1}) - K_i \log(K_l/K_{l-1})] \cdot \mathbb{1}(K_i < K_l)$, $l = 1, 2, \ldots, q+1$ for call options. From **Eq. 2**, a_{q+1} can be represented by $a_1, a_2, \ldots, a_q$, as

$$a_{q+1} = \left(1 - \sum_{l=1}^{q} a_l \log \frac{K_l}{K_{l-1}}\right) (\log c_K)^{-1} \tag{S3}$$

Plugging **Eq.** S3 into **Eqs** S1, S2, we obtain

$$\begin{aligned} e^{R_t T}\hat{P}_i &= \sum_{l=1}^{q+1} a_l X_{i,l}^{(p)} \\ &= a_1 X_{i,1}^{(p)} + a_2 X_{i,2}^{(p)} + \cdots + a_q X_{i,q}^{(p)} \\ &\quad + \left(1 - a_1 \log\frac{K_1}{K_0} - \cdots - a_q \log\frac{K_q}{K_{q-1}}\right)(\log c_K)^{-1} X_{i,q+1}^{(p)} \\ &= a_1[X_{i,1}^{(p)} - (\log\frac{K_1}{K_0})(\log c_K)^{-1} X_{i,q+1}^{(p)}] \quad + \cdots \\ &\quad + a_q[X_{i,q}^{(p)} - (\log\frac{K_q}{K_{q-1}})(\log c_K)^{-1} X_{i,q+1}^{(p)}] + \frac{1}{\log c_K} X_{i,q+1}^{(p)} \\ &\triangleq a_1 X_{i,1}^{(P)} + a_2 X_{i,2}^{(P)} + \cdots + a_q X_{i,q}^{(P)} + X_{i,q+1}^{(P)},\ i \in \mathcal{P} \end{aligned} \tag{S4}$$

where $X_{i,l}^{(P)} = X_{i,l}^{(p)} - (\log K_l/K_{l-1})(\log c_K)^{-1} X_{i,q+1}^{(p)}$, $l = 1, 2, \ldots, q$ and $X_{i,q+1}^{(P)} = X_{i,q+1}^{(p)}/\log c_K$. Similarly for call options,

$$\begin{aligned} e^{R_t T}\hat{C}_i &= \sum_{l=1}^{q+1} a_l X_{i,l}^{(c)} \\ &= a_1 X_{i,1}^{(c)} + a_2 X_{i,2}^{(c)} + \cdots + a_q X_{i,q}^{(c)} \\ &\quad + \left(1 - a_1 \log\frac{K_1}{K_0} - \cdots - a_q \log\frac{K_q}{K_{q-1}}\right)(\log c_K)^{-1} X_{i,q+1}^{(c)} \\ &= a_1[X_{i,1}^{(c)} - (\log\frac{K_1}{K_0})(\log c_K)^{-1} X_{i,q+1}^{(c)}] \quad + \cdots \\ &\quad + a_q[X_{i,q}^{(c)} - (\log\frac{K_q}{K_{q-1}})(\log c_K)^{-1} X_{i,q+1}^{(c)}] + \frac{1}{\log c_K} X_{i,q+1}^{(c)} \\ &\triangleq a_1 X_{i,1}^{(C)} + a_2 X_{i,2}^{(C)} + \cdots + a_q X_{i,q}^{(C)} + X_{i,q+1}^{(C)},\ i \in \mathcal{C} \end{aligned} \tag{S5}$$

where $X_{i,l}^{(C)} = X_{i,l}^{(c)} - (\log K_l/K_{l-1})(\log c_K)^{-1} X_{i,q+1}^{(c)}$, $l = 1, \ldots, q$ and $X_{i,q+1}^{(C)} = X_{i,q+1}^{(c)}/\log c_K$. □

B. PROOF OF THEOREM 3.1

Given $\epsilon > 0$, let $\delta_1 = \sqrt{\epsilon} e^{R_t T}/[3(1 + c_K + e)] > 0$. There exists $-\infty < A < 0 < B < \infty$, such that,

$$\int_{-\infty}^{A} f_{\mathbb{Q}}(x)dx < \delta_1,\ \int_{-\infty}^{A} e^x f_{\mathbb{Q}}(x)dx < \delta_1,\ \int_{B}^{\infty} f_{\mathbb{Q}}(x)dx < \delta_1,\ \int_{B}^{\infty} e^x f_{\mathbb{Q}}(x)dx < \delta_1$$

Let $\delta_2 = \sqrt{\epsilon} e^{R_t T - B - 1}/[3(B - A + 2)] > 0$. Since $f_{\mathbb{Q}}$ is continuous, there exists a $\delta > 0$, such that, for any $x_1, x_2 \in [A - 1, B + 1]$,

$$|f_{\mathbb{Q}}(x_1) - f_{\mathbb{Q}}(x_2)| < \delta_2$$

as long as $|x_1 - x_2| < \delta$.

For small enough $K_1, |\Delta|$ and large enough q, K_q, there exist integers u, v, such that, $1 < u < u+1 < v < v+1 < q$, $\log K_u \leq A < \log K_{u+1}$, $\log K_v < B \leq \log K_{v+1}$, $|\Delta| < \delta$.

We construct a f_Δ by defining

$$\begin{aligned} a_1 &= (\log c_K)^{-1} \int_{-\infty}^{\log K_1} f_{\mathbb{Q}}(x)dx \ \geq \ 0 \\ a_i &= [\log(K_i/K_{i-1})]^{-1} \int_{\log K_{i-1}}^{\log K_i} f_{\mathbb{Q}}(x)dx \ \geq \ 0, \quad i = 2, \ldots, q \\ a_{q+1} &= (\log c_K)^{-1} \int_{\log K_q}^{\infty} f_{\mathbb{Q}}(x)dx \ \geq \ 0 \end{aligned}$$

It can be verified that $\int_{-\infty}^{\infty} f_\Delta(x)dx = \sum_{i=1}^{q+1} a_i \log(K_i/K_{i-1}) = 1$. Let

$$\Delta_f = \max_{u \leq i \leq v} \left(\max_{\log K_i \leq x \leq \log K_{i+1}} f_{\mathbb{Q}}(x) - \min_{\log K_i \leq x \leq \log K_{i+1}} f_{\mathbb{Q}}(x) \right)$$

Then $|\Delta| < \delta$ implies $\Delta_f \leq \delta_2$. It can be verified that

$$\begin{aligned} |\hat{C}_i - \tilde{C}_i| &< \begin{cases} \sqrt{\epsilon}/3, & \text{for } i = v+1, \ldots, q \\ 2\sqrt{\epsilon}/3, & \text{for } i = u, \ldots, v \\ \sqrt{\epsilon}, & \text{for } i = 1, \ldots, u-1 \end{cases} \\ |\hat{P}_i - \tilde{P}_i| &< \begin{cases} \sqrt{\epsilon}/3, & \text{for } i = 1, \ldots, u \\ 2\sqrt{\epsilon}/3, & \text{for } i = u+1, \ldots, v+1 \\ \sqrt{\epsilon}, & \text{for } i = v+2, \ldots, q \end{cases} \end{aligned}$$

In other words, there exist $a_1, \ldots, a_{q+1}$, such that, $(\hat{C}_i - \tilde{C}_i)^2 < \epsilon$, $(\hat{P}_i - \tilde{P}_i)^2 < \epsilon$, for $i = 1, \ldots, q$. It implies the $(a_1, \ldots, a_{q+1})$ that minimizes $L(a_1, \ldots, a_{q+1})$ also satisfies

$$\frac{1}{2q} \left[\sum_{i=1}^{q} (\hat{C}_i - \tilde{C}_i)^2 + \sum_{i=1}^{q} (\hat{P}_i - \tilde{P}_i)^2 \right] < \epsilon$$

which leads to the conclusion. □

C. PROOF OF PROPOSITION 4.1

Since $\mathbb{E}_t^{\mathbb{Q}}[\sum_{i=1}^{T} R_i^2] = \sum_{i=1}^{t} R_i^2 + \sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[R_i^2]$, the key part

$$
\begin{aligned}
&\sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[R_i^2] = \sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[\log \frac{S_i}{S_{i-1}}]^2 \\
&= \sum_{i=t+1}^{T} [\mathbb{E}_t^{\mathbb{Q}}(\log S_i)^2 + \mathbb{E}_t^{\mathbb{Q}}(\log S_{i-1})^2 - 2\mathbb{E}_t^{\mathbb{Q}}(\log S_i)(\log S_{i-1})] \\
&= \sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}(\log S_i)^2 + \sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}(\log S_{i-1})^2 - 2\sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[\log S_{i-1} + \log(\frac{S_i}{S_{i-1}})][\log S_{i-1}] \\
&= \sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}(\log S_i)^2 + \sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}(\log S_{i-1})^2 - 2\sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}(\log S_{i-1})^2 \\
&\quad - 2\sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[\log S_{i-1}][\log(\frac{S_i}{S_{i-1}})] \\
&= \mathbb{E}_t^{\mathbb{Q}}[\log S_T]^2 - [\log S_t]^2 - 2\sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[\log S_{i-1}][\log(\frac{S_i}{S_{i-1}})] \\
&= \mathbb{E}_t^{\mathbb{Q}}[\log S_T]^2 - [\log S_t]^2 - 2\sum_{i=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[\log S_{i-1}]\mathbb{E}_t^{\mathbb{Q}}[\log(\frac{S_i}{S_{i-1}})] \\
&= \mathbb{E}_t^{\mathbb{Q}}[\log S_T]^2 - [\log S_t]^2 - 2\sum_{i=t+1}^{T} [\mathbb{E}_t^{\mathbb{Q}} \log S_{i-1}\mathbb{E}_t^{\mathbb{Q}} \log S_i - (\mathbb{E}_t^{\mathbb{Q}} \log S_{i-1})^2]
\end{aligned}
$$

Then **Eq. 12** can be obtained by plugging $\mathbb{E}_t^{\mathbb{Q}}[\sum_{i=1}^{T} R_i^2]$ into **Eq. 11**. □

D. LINEAR INTERPOLATION FOR 1ST AND 2ND MOMENTS IN SECTION 4.1

Mean imputation Suppose the trading day is t and the expiration day is T. We denote all possible expiration dates of traded contracts by $t + n_1, t + n_2, \ldots$. Suppose the time point to be imputed is $t + n_0$. Given all the information available at day t, $\log S_t$ can be regarded as its expectation at day t, $\mathbb{E}_t^{\mathbb{Q}} \log S_t$. Therefore, we consider cases separately according to whether or not $t + n_0$ is in the interval $[t, t + n_1]$ and then apply linear interpolation to obtain the mean of $\log S_{t+n_0}$. More specifically, there are two cases:

Case 1: $n_0 \in [0, n_1]$ and $\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_1})$ has been calculated.

$$
\begin{aligned}
\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_0}) &= \mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_1}) - \frac{(n_1 - n_0)[\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_1}) - \log S_t]}{n_1} \\
&= \frac{n_0\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_1}) + (n_1 - n_0)\log(S_t)}{n_1}
\end{aligned}
$$

Case 2: $n_0 \in [n_i, n_{i+1}]$ for some $i = 1, 2, \ldots$. The expectations $\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_i})$ and $\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_{i+1}})$ have already been calculated.

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_0}) &= \frac{(n_0 - n_i)[\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_{i+1}}) - \mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_i})]}{n_{i+1} - n_i} + \mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_i}) \\ &= \frac{(n_0 - n_i)\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_{i+1}}) + (n_{i+1} - n_0)\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_i})}{n_{i+1} - n_i}\end{aligned}$$

Variance Imputation In order to calculate the variance $\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_0})$ at day t, we use a similar interpolation based on the available variances of log returns at day t with expiration T. Based on the scatterplot (not shown here) of all available variances that we have from the existing contracts, the trend of variances has a curved pattern against the number of days to expiration. More specifically, it is roughly a quadratic curve. Before we implement a linear interpolation, we first perform a square-root transformation of variances.

Case 1: $n_0 \in [0, n_1]$. $\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_1})$ has been calculated. Then

$$\sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_0})} = \frac{n_0\sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_1})}}{n_1}$$

Case 2: $n_0 \in [n_i, n_{i+1}]$ for some $i = 1, 2, \ldots$. The values $\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_i})$ and $\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_{i+1}})$ have been calculated. Then

$$\begin{aligned}&\sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_0})} \\ &= \sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_0})} - \sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_i})} + \sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_i})} \\ &= \frac{(n_0 - n_i)\left[\sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_{i+1}})} - \sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_i})}\right]}{n_{i+1} - n_i} + \sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_i})} \\ &= \frac{(n_0 - n_i)\sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_{i+1}})} + (n_{i+1} - n_0)\sqrt{\mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_i})}}{n_{i+1} - n_i}.\end{aligned}$$

Then the second moment is

$$\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_0})^2 = [\mathbb{E}_t^{\mathbb{Q}}(\log S_{t+n_0})]^2 + \mathbb{V}_t^{\mathbb{Q}}(\log S_{t+n_0})$$

A fair price of variance swap $VS_{t,T}$ can be obtained by the pricing formula **Eq. 11**.